

Elizabeth Reis Teixeira

# O sistema segmental do português

constituição e  
simplificação fonológica



Este trabalho nasceu da compilação de textos, criados para as aulas de Fonética e Fonologia do Português do Instituto de Letras da Universidade Federal da Bahia (UFBA), na década de 1980, com o objetivo de apresentar os aspectos fonoarticulatórios constitutivos e fonológicos distintivos – tanto no que diz respeito ao eixo de composição como ao eixo de combinação – do sistema de sons do português.

A introdução dessa base fonoarticulatória e do arcabouço silábico-lexical do português brasileiro subsidia nossas análises sobre a constituição e simplificação dos sistemas fonológicos infantis, levando às principais contribuições de nossas pesquisas para a área de estudos da Aquisição da Linguagem tanto no que diz respeito ao desenvolvimento normal como em relação a algumas condições atípicas, como a dislalia, a dislexia, e o atraso de linguagem na deficiência intelectual.

Como decorrência, estudos de frequência de ocorrência dos padrões segmentais, silábicos, lexicais, intrassilábicos e intersilábicos mais afetados tanto na fala infantil como na fala adulta no português foram levantados estatisticamente, bem como padrões de simplificação indicadores de diferentes graus de estigmatização sociolinguística que podem influenciar as tendências aquisicionais.

# **O sistema segmental do português**

constituição e simplificação fonológica



UNIVERSIDADE FEDERAL DA BAHIA

Reitor

*Paulo Cesar Miguez de Oliveira*

Vice-reitor

*Penildon Silva Filho*



EDITORA DA UNIVERSIDADE FEDERAL DA BAHIA

Diretora

*Susane Santos Barros*

Conselho Editorial

Titulares

*Angelo Szaniecki Perret Serpa*

*Caiuby Alves da Costa*

*Cleise Furtado Mendes*

*George Mascarenhas de Oliveira*

*Mônica de Oliveira Nunes de Torrenté*

*Mônica Neves Aguiar da Silva*

Suplentes

*José Amarante Santos Sobrinho*

*Lorene Pinto*

*Lúcia Matos*

*Lynn Alves*

*Paola Berenstein Jacques*

*Rafael Moreira Siqueira*

Elizabeth Reis Teixeira

# **O sistema segmental do português**

constituição e simplificação fonológica

Salvador  
Edufba  
2025

2025, Elizabeth Reis Teixeira.  
Direitos para esta edição cedidos à Edufba. Feito o Depósito Legal.  
Grafia atualizada conforme o Acordo Ortográfico da Língua Portuguesa de 1990,  
em vigor no Brasil desde 2009.

Coordenação editorial  
*Cristovão Mascarenhas*

Coordenação gráfica  
*Edson Nascimento Sales*

Coordenação de produção  
*Gabriela Nascimento*

Assistente editorial  
*Aline Silva Santos*

Capa e projeto gráfico  
*Rodrigo Oyarzábal Schlabit*

Revisão  
*Clara Moraes*

Normalização  
*Marcelly Moreira Santos*

Imagem da capa  
*Elizabeth Reis Teixeira, via ChatGPT*

Sistema Universitário de Bibliotecas – SIBI/UFBA

---

Teixeira, Elizabeth Reis.

O sistema segmental do português : constituição e simplificação fonológica /  
Elizabeth Reis Teixeira. - Salvador : EDUFBA, 2025.  
179 p.

ISBN 978-65-5630-829-6

1. Língua portuguesa - Fonologia. 2. Fala. 3. Linguística. 4. Aquisição de linguagem.  
I. Título.

CDD 469.15

---

Elaborada por Selma Matos / CRB-5: BA-1001

Editora afiliada à



Editora da UFBA

Rua Barão de Jeremoabo, s/n – *Campus* de Ondina  
40170-115 – Salvador, Bahia / Tel.: +55 71 3283-6164  
edufba@ufba.br / edufba.ufba.br/

Falamos uma média de quinze consoantes e vogais por segundo, e conseguimos organizar de forma ordenada estes enunciados em unidades maiores de fala chamadas ‘sílabas’ ao circundar nossas vogais com consoantes de diversas maneiras.

Na fala, escutamos uma consequência acústica que representa a soma dos movimentos de uma dada consoante ou vogal. Portanto, cada padrão de quinze ou mais ações musculares se reduz a um som. Consequentemente, a incrível versatilidade do sistema de ações da fala, que se coloca como único no reino animal, não consegue ter o respeito que merece, nem da ciência nem do mundo em geral. É, na verdade, um milagre invisível<sup>1</sup> (MacNeilage, 2010, p. 4-5, tradução nossa).

---

1 “We speak at the rate of some fifteen consonants and vowels per second, and we manage to neatly organize these utterances into larger output chunks called ‘syllables’ by surrounding our vowels with consonants in various ways. In speech we just hear one acoustical consequence that represent the sum of the movements for a given consonant or vowel. Thus, every pattern of fifteen or so muscle actions boils down to one sound. Consequently, the astounding versatility of the speech action system, which is a league of its own in the animal kingdom, doesn’t begin to get the respect it deserves, either in science or in the world in general. It is, in effect, an invisible miracle”.





# SUMÁRIO

9	<b>Apresentação</b>
15	<b>Introdução</b>
21	<b>Capítulo 1</b> <b><i>Bases fonoarticulatórias</i></b>
39	<b>Capítulo 2</b> <b><i>O sistema fonológico do português: composição e combinação</i></b>
69	<b>Capítulo 3</b> <b><i>Simplificação fonológica: aquisição da fala</i></b>
91	<b>Capítulo 4</b> <b><i>Simplificação fonológica em situações aquisicionais atípicas</i></b>
113	<b>Capítulo 5</b> <b><i>Simplificação fonológica na aquisição da escrita inicial</i></b>
119	<b>Capítulo 6</b> <b><i>Frequência da ocorrência dos padrões segmentais, silábicos, lexicais, intrassilábicos e intersilábicos no português</i></b>

141    **Capítulo 7**

*Processos de simplificação fonológica como marcadores de  
estigmatização sociolinguística*

149    **Capítulo 8**

*Erros de ordenação serial e permutação segmental intersilábica*

163    **Considerações finais**

171    **Referências**

# Apresentação

O presente trabalho nasceu da compilação de textos por nós criados para serem utilizados nas aulas de Fonética e Fonologia do Português do Instituto de Letras da Universidade Federal da Bahia (UFBA), a partir de meados da década de 1980, na disciplina “Língua Portuguesa IV – Fonologia do Português (LET 369)”, do Curso de Graduação em Letras – hoje “Fonética e Fonologia da Língua Portuguesa (LET A17)” –, e na disciplina do Curso de Pós-Graduação “Fonologia do Português (LET 677)”.

Sua primeira sistematização, intitulada “Módulo I: Noções Básicas de Linguística”, ocorreu em 1993, quando impresso para utilização no Curso de Atualização em Língua Portuguesa para Professores de 2º grau, organizado, conjuntamente, pela Fundação Vitae, Secretaria de Educação do Estado da Bahia e UFBA. Passou por revisões em 1998, 2006 e foi revisto e aumentado para ser disponibilizado, em 13 de janeiro de 2010, no espaço virtual da UFBA – Plataforma Moodle – versão 1.6, no Ambiente Virtual de Aprendizagem (AVA) da disciplina LET A17, com o título de “Aspectos Fonoarticulatórios e Fonológicos do Português”.

Nessas primeiras versões, apresentavam-se, com a finalidade de introduzir o estudo do sistema de sons do português, os aspectos fonoarticulatórios constitutivos (tanto no que diz respeito ao eixo de composição como ao eixo de combinação) dos diversos elementos vocálicos e consonantais que constituem o inventário de sons da língua. Em um segundo momento, listavam-se os elementos distintivos do sistema fonológico da

língua a partir de uma concepção de *análise polissistêmica*, em conformidade à Teoria Prosódica de Firth (1957), que considera a necessidade de estabelecer sistemas independentes e não relacionados em posições distintas da estrutura (como as posições inicial e final da sílaba), em vez de apenas um sistema uniforme de elementos fonológicos que se instanciam em todos os ambientes (muito embora sujeitos a restrições de distribuição). Isso significa dizer que, se a abrangência dos contrastes entre os elementos funcionais que aparecem em duas posições distintas não é a mesma, estes não podem ser considerados idênticos, mesmo que tenham a mesma natureza fonética. Como bem colocado por Anderson (1985, p. 182, tradução nossa):

[...] para Firth não fazia sentido falar sobre um único sistema de elementos subjacentes a uma língua, que de alguma forma é tarefa do linguista descobrir. As estruturas e elementos não estão, de forma alguma, presentes em uma língua independentemente da análise do linguista: eles são apenas abstrações que o linguista faz sobre os fenômenos de uso da língua, e sua meta é fornecer uma estrutura conceitual para entender a língua, em vez de apresentar uma estrutura que tenha status ontológico independente<sup>2</sup>.

A adoção desse tipo de análise polissistêmica justifica-se pela facilidade que oferece ao tratamento dos dados de aquisição fonológica – em que o sistema adulto alvo, através de padrões de simplificação que a criança vai descartando, acaba se constituindo – e para os quais dedicamos, quase que inteiramente, nossos interesses e nossas atividades de pesquisa.

Na versão atual, decidimos, portanto, incluir os estudos que motivaram a real necessidade de nossa análise fonológica da língua: os estudos de aquisição da fonologia do português em condições normais e atípicas (especificamente, os casos de desvios fonológicos evolutivos), iniciados a

---

2 “[...] for Firth it was meaningless to speak of a single system of elements underlying a language, which in some sense is the linguist's task to discover. Structures and elements are not in any way present in a language independent of the linguist's analysis: they are merely abstractions the linguist makes from the phenomena of language use, and their goal is to provide a conceptual structure for understanding language use rather than to present some structure which has independent ontological status”.

partir das pesquisas desenvolvidas quando estudante do Departamento de Fonética do University College, da Universidade de Londres (Teixeira, 1980, 1985).

Achamos que seria uma decorrência natural incluirmos, também, os estudos sobre os padrões de simplificação do sistema fonológico da língua que se estenderam para pesquisas sobre outras condições de desenvolvimento atípico, através de trabalhos de mestrado e doutorado sob nossa orientação (Iácono, 2014; Melo, 2011; Mello, 2024; Pepe, 2010) sobre a aquisição da escrita inicial (Almeida, 2013); bem como estudos sobre a recorrência de padrões fonológicos na língua (Teixeira, 1997; Teixeira; Davis, 2002; Teixeira; Silva, 1999; Silveira, 2003, 2005); pesquisas sobre erros de ordenação serial (Teixeira, 2003a, 2003b); e sobre a simplificação fonológica como marca de estigmatização sociolinguística (Teixeira, 1986).

Seguindo a estrutura das versões anteriores, a “Introdução” apresenta, do ponto de vista psicolinguístico, as questões que giram em torno da especialização do homem para a fala e sua aquisição, e a importância dos padrões organizados de sons na constituição dos sistemas de comunicação humana.

O primeiro capítulo introduz as “Bases fonoarticulatórias”, ou seja, fornece os pressupostos básicos em termos da fonética articulatória para a descrição e classificação dos sons produzidos nas línguas naturais. *Articulações consonantais*, *Articulações vocálicas*, *Ditongos*, a *Nasalização*, a *Estrutura da sílaba* e a *Estrutura da palavra* são introduzidos como pressupostos teóricos para a identificação dos elementos distintivos da língua a serem inventariados no capítulo seguinte.

O Capítulo 2, “O sistema fonológico do português: composição e combinação”, apresenta o inventário de elementos distintivos do sistema fonológico da língua, a partir de uma visão polissistêmica, *i.e.*, apresentando os segmentos em termos de sua ocorrência em diferentes posições da estrutura silábico-prosódico-lexical: *Margem inicial* e *Margem final* da sílaba no que diz respeito ao sistema consonantal; e *Posição tônica*, *Posição pré-tônica*, *Posição pós-tônica não final* e *Posição pós-tônica final* para as vogais. *Ditongação (Sequências vocálicas)*, *Nasalização* e *Estrutura silábica* também são tratadas. Tendências fonológicas e possibilidades realizacionais são discutidas em todos os casos. A análise polissistêmica

justifica-se devido a sua propriedade para o tratamento e descrição dos padrões de simplificação fonológica encontrados na aquisição – que serão amplamente discutidos no capítulo que se segue.

No Capítulo 3, apreciamos os padrões gerais de simplificação que afetam as classes de sons e suas possibilidades combinatórias durante a aquisição fonológica em português, com o intuito de mostrar como, pouco a pouco, a criança vai construindo e constituindo o seu sistema fonológico, espelhada no modelo e alvo adulto. Mostramos, para tanto, o arcabouço teórico que propomos para a análise dos sistemas fonológicos em desenvolvimento, com base nos pressupostos da Fonologia Natural – uma tipologia detalhada para classificação dos padrões de simplificação em termos de: *Processos de substituição*, *Processos modificadores estruturais* e *Processos sensíveis ao contexto*.

O Capítulo 4, “Simplificação fonológica em situações aquisicionais atípicas”, mostra como o arcabouço teórico de *Processos de simplificação fonológica*, desenvolvido para a análise da fala de crianças adquirindo o português, foi aplicado para o estudo da aquisição em condições linguísticas não esperadas, ou não normais. São examinados, aqui, casos de simplificação fonológica nos desvios fonológicos de desenvolvimento (ou dislalia), na dislexia, na deficiência intelectual e no Transtorno do Espectro Autista (TEA).

O Capítulo 5, por sua vez, demonstra como os padrões de simplificação fonológica podem ser, também, empregados para o estudo da aquisição da escrita inicial em português. Os *Processos modificadores estruturais* ocorrem com maior frequência, seguidos, em menor grau, pelos de *Substituição*. Os *Processos sensíveis ao contexto*, tão recorrentes nos estágios iniciais da aquisição da fala, não apresentam, aqui, ocorrência significativa.

No Capítulo 6, apresentamos os estudos sobre contagem de frequência de ocorrência dos padrões segmentais, silábicos, lexicais, intrassilábicos e intersilábicos no português como uma forma de mostrar, estatisticamente, os padrões constitutivos do sistema fonológico que tem maior representatividade na língua, e que vem a ser mais simplificados tanto na fala infantil como na fala adulta.

O Capítulo 7 aborda a questão dos processos de simplificação fonológica que têm sido reportados na literatura como marcadores de estigmatização sociolinguística: *Simplificação dos ditongos crescentes*, *Simplificação*

*das consoantes finais* e *Simplificação dos encontros consonantais*. Esses padrões de simplificação, presentes na fala das populações com baixa ou nula escolaridade, têm atuado como indicadores de diferentes graus de estigmatização social.

No Capítulo 8, apresentamos os *Erros de ordenação serial* que ocorrem na fala adulta e o correlato processo de *Permutação segmental intersilábica*, encontrado como um padrão simplificacional na fala infantil. O estudo dos erros de ordenação serial revela a importância da estrutura do molde silábico como fator primordial na constituição dos sistemas fonológicos tanto do ponto de vista ontogenético como filogenético.

O capítulo final expõe, nas “Considerações finais”, as contribuições que os estudos sobre a constituição e simplificação dos sistemas fonológicos trazem ao cenário atual da Aquisição da Linguagem, particularmente à área da Aquisição Fonológica, e apresenta sugestões para estudos futuros.





# Introdução

A linguagem foi, sem dúvida, a aquisição mais marcante em termos do desenvolvimento da espécie humana. Embora outras espécies no reino animal tenham conseguido desenvolver sistemas de comunicação bastante efetivos – como as abelhas e os golfinhos –, a capacidade de abstração e a infinita gama de possibilidades de combinação de seus elementos distinguem a linguagem humana da comunicação de outras espécies no reino animal. Ao abstrair, podemos não só falar de coisas que estão ocorrendo no momento do ato de fala, mas também nos reportar a eventos passados e a situações que ocorrerão no futuro.

As teorias sobre a origem e o desenvolvimento da fala no homem remontam aos relatos bíblicos sobre a origem divina da linguagem. Alguns teóricos sugerem que a fala foi desenvolvida bastante recentemente: há 40 mil anos, durante o quarto período glacial; enquanto outros propõem que ela teria se desenvolvido entre 2 e 3 milhões de anos atrás (Bickerton, 1990).

Não há nenhum campo de investigação que possa fornecer, isoladamente, evidência suficiente, mas, se examinarmos os achados provenientes do estudo de fósseis, dos estudos sobre a comunicação vocal e gestual em criaturas vivas e dos estudos sobre o cérebro e o aparelho fonador (o trato vocal), podemos nos aproximar de uma teoria sobre como a linguagem e a fala se desenvolveram, e até mesmo sugerir como estas se originaram.

Uma questão permanece, contudo, bastante intrigante: o desenvolvimento de um complexo sistema de comunicação entre homínídeos só

foi possível por ter sido a fala o seu principal modo de transmissão? Ao que tudo indica, a fala mostrou-se o meio mais vantajoso por uma série de razões. Primeiramente, o canal auditivo-vocal facilitava a comunicação em condições em que um outro canal falharia, como seria o caso do gestual-visual. A utilização da fala permitia que a comunicação ocorresse, por exemplo, ao mesmo tempo em que tarefas manuais estivessem sendo executadas, tais como a manufatura de armas ou a colheita de alimentos. Em segundo lugar, a fala exibe certas características acústicas, tais como a *coarticulação* e a *redundância*, que a tornam um meio supereficiente de comunicação. A coarticulação permite que os falantes produzam sons em paralelo, integrando as características acústicas e articulatórias de um segmento a de outros segmentos adjacentes. Um bom exemplo disso é a pronúncia do fonema /j/ quando seguido da vogal /u/, como ocorre na primeira sílaba da palavra “CHUVA”: o arredondamento e a protrusão labiais, típicos do /u/, já se fazem presentes na produção da consoante fricativa. Esses fatos de coarticulação resultam na transmissão da informação fonética e linguística de forma extremamente rápida. O alto grau de redundância, por sua vez, se explica devido às diversas pistas contidas na produção de um único segmento. Voltando ao exemplo da primeira consoante da palavra CHUVA, o /j/, as pistas em relação ao modo de articulação são fornecidas não só pelo estreitamento do canal oral à passagem do ar na região palato-alveolar (ou pós-alveolar), mas também na região labial (*i.e.*, existe uma maior aproximação dos dois lábios na realização da consoante devido à preparação dos órgãos articulatórios para a produção da vogal [u] subsequente). Além disso, existem, ainda, os padrões entoacionais, de acentuação, de junturas entre palavras e frases e outras características paralinguísticas, tais como o tom vocal e as pistas faciais, que ajudam a estabelecer informações sintáticas, semânticas e fonológicas, impedindo que a mensagem possa ser, eventualmente, distorcida e que haja alguma falha que afete a inteligibilidade.

Segundo MacNeilage (2010), o fato de que durante cada segundo de fala, inconscientemente, usamos quase 225 ações musculares torna o sistema da fala humana único no reino animal, e sua aquisição torna pequena a maior parte das conquistas evolutivas de outras espécies. Seguindo uma abordagem neodarwiniana da fala, o autor propõe que, através de um

processo de descendência, as capacidades vocais ancestrais modificaram-se como resposta às pressões da seleção natural em busca de maior eficiência na comunicação. A linguagem desenvolveu-se, primeiramente, na modalidade vocal auditiva, porque sua estrutura prototípica possibilitou uma solução mais natural ao problema da expansão do léxico do que as possibilidades que eram oferecidas pelo sistema manual. A alternância aberta-fechada, que forneceu o molde para as sílabas, desenvolveu-se a partir de ciclicidades ingestivas tais como a mastigação, através de um estágio intermediário de ciclicidades visuofaciais como os estalos labiais, que são comuns em outros primatas superiores.

A observação sobre o processo pelo qual as crianças adquirem uma língua revela as limitações que já estavam presentes em nossos ancestrais quando estes começaram a falar, ou seja, nesse sentido, a ontogênese parece recuperar a filogênese.

Moldes motores para a fala evoluíram de ciclicidades mandibulares via um estágio intermediário de estalos comunicativos visuofaciais, que eventualmente se acoplaram à fonação para formar as proto-sílabas. Estas proto-sílabas, primeiro, tiveram um papel de arrumação vocal. Elas tornaram-se efetivas porque permitiram a transmissão omnidirecional de um sinal comunicativo padrão, prontamente estendido ao longo do tempo, e com alternâncias acústicas agudas entre estados abertos e fechados de complexidade suficiente para segurar o interesse do ouvinte. Em algum ponto, um dos conjuntos limitados de formas proto-silábicas – uma variante nasalizada – pareceu-se ao conceito parental de fêmea, resultando na forma [mama].

Esta foi uma invenção social e momentânea. Abriu o caminho para uma série de eventos singulares semelhantes que ligaram, um por um, itens adicionais a dois, até então, níveis desvinculados de função – conceito e padrões sonoros.

Uma vez que esta invenção passou da matriz infante-pais para a sociedade em geral, pareamentos subsequentes de conceito/som para novas palavras estabeleceram-se por concordância cultural, e a história desses acordos foi passada adiante a gerações sucessivas de usuários da língua (MacNeilage, 2010, p. 236).

A propriedade bifásica do ciclo mandibular (o molde) forneceu a base inicial para a estrutura segmentável na forma do primeiro corte binário no domínio do tempo. Os dois movimentos opostos do ciclo fornecem, por definição, dois estágios distintos, livres e aproximados: a boca aberta, acompanhada pela fonação, resulta em um *output* acústico de alta amplitude; a boca fechada – e a boca totalmente fechada é considerada a forma modal inicial – resulta em uma queda drástica (*damping*) da fonte de sonoridade e, conseqüentemente, um *output* acústico de baixa amplitude.

Considerando que as primeiras palavras tenham sido termos de família como “MAMA” e “PAPA”, observa-se que o primeiro é produzido por interações do molde puro, *i.e.*, dois ciclos de oscilação mandibular sem movimento da língua: existem dois picos de amplitude ao sinal correspondente às duas vogais e dois períodos de frequência relativamente baixa e atividade de baixa amplitude associados às duas oclusões dos lábios na produção dos dois “m”s. Concomitantemente, o véu palatino está rebaixado para abrir a passagem à cavidade nasal. A forma paterna “PAPA” envolve os mesmos dois primeiros ciclos do molde puro, e a passagem do véu palatino à cavidade nasal permanece fechada durante a produção de todo o enunciado.

Portanto, nesse pequeno microcosmo de duas palavras, temos a base da estrutura segmentável – a possibilidade de variação independente de consoantes e vogais no domínio do tempo. Duas consoantes podem ocorrer como resultado de apenas uma diferença articulatória com, basicamente, a mesma vogal. Este é um segundo estágio que vai além do molde com a fonação própria – na direção de uma solução específica para produção da fala em relação à questão do ordenamento serial. Permite a produção de sequências diferentes dos mesmos elementos. Erros de ordenação serial, que tratamos em capítulo adiante, mostram que o adulto moderno tem representações separadas para consoantes e vogais, *i.e.*, erros de fala são subsilábicos, envolvendo elementos ordenados em série dentro da sílaba: as permutações de elementos entre sílabas ocorrem de forma independente entre consoantes e vogais, respeitando as estruturas tanto silábica como lexical, *i.e.*, consoantes permutam apenas com consoantes e vogais permutam apenas com vogais.

Ex.: MUSSARELA DE BÚFALA [busa'relə dʒi 'mufalə]

MILHO E ERVILHA [miʔə i eɣ'viʔu]

Nossos sistemas articulatório, fonatório e respiratório não evoluíram apenas para servir às funções de comer e produzir ruídos, mas para servir às funções específicas da produção da fala articulada. É exatamente a ação integrada desses mecanismos que torna a manifestação físico-acústica da fala o mais efetivo meio de trocas linguísticas. Isso quer dizer que, psicolinguisticamente, um indivíduo falante “normal” (*i.e.*, sem comprometimento de ordem central ou periférica):

- tem total controle sobre sua capacidade de fala;
- emite sons como resposta reflexiva a emoções vivenciadas (como, por exemplo, uma gargalhada), mas consegue, sobretudo, produzir vocalizações voluntárias como expressão de eventos fora do interlocutor e que transcendem o seu *aqui* e *agora* (isto é, como resposta a estímulos arbitrários);
- consegue concatenar, de inúmeras formas, as unidades que compõem o sistema fonológico da língua em unidades significativas maiores (como a palavra, a frase, a oração, ou mesmo o texto), num eixo de sucessividade e de acordo com esquemas de processamento e organização altamente sofisticados.

Portanto, embora tenhamos em comum com outros animais certas estruturas anatômicas, como boca, língua, lábios, maxilares providos de dentes e a habilidade de utilizar a respiração para produzir sons que ressoam nas cavidades oral e nasal, tendo em vista as características especiais de nosso cérebro e de nosso aparato fonoarticulatório, a *linguagem* humana, que se manifesta através de qualquer língua natural em forma de *fala* (exceto no caso das línguas sinalizadas, naturais aos surdos), nada mais é do que uma capacidade que tem uma base biológica específica a nossa espécie. Isso quer dizer que, para pôr em ação sua capacidade linguística, todo indivíduo necessita dominar uma língua, um código linguístico qualquer, que lhe permita interagir com a comunidade linguística na qual se encontra inserido. Uma *língua natural*, portanto, é toda aquela adquirida naturalmente, *i.e.*, sem necessidade de instrução e intervenção formal e sistemática, mas através, apenas, da exposição do indivíduo a um meio linguístico específico (ou mais de um, no caso de comunidades não monolíngues) em que essa língua ocupe papel preponderante.

Nos capítulos que se seguem, apresentamos informações sobre as características fonoarticulatórias dos sistemas de sons das línguas naturais, descrevendo como os sons que constituem os sistemas linguísticos dessas línguas orais são produzidos pelo aparato fonador humano.



## Capítulo 1

# Bases fonoarticulatórias

Os três mecanismos básicos necessários para a produção de cadeias sonoras, na fala, são:

- o mecanismo respiratório;
- o mecanismo fonatório;
- e o mecanismo articulatório.

A fonte primordial de energia para a produção da maioria dos sons da fala é o movimento respiratório que expulsa o ar dos pulmões. O ar proveniente dos pulmões – extraído pela compressão dos músculos torácicos e abdominais – sobe pela *traqueia* até a *laringe*, onde passa por um estreito canal compreendido entre duas dobras musculares, comumente chamadas de *cordas vocais*. O espaço compreendido entre esse par de ligamentos é chamado de *glote*.

Se as cordas vocais estão separadas (*i.e.*, uma dobra vocal distante da outra), como ocorre na expiração normal, o ar que vem dos pulmões passa livremente e segue seu caminho pelo canal da faringe até atingir a cavidade oral e, finalmente, escapar pela boca. Sons produzidos com as cordas vocais nessa posição (ou seja, sem a presença de excitação ou energia vocal) são chamados de *surdos*. Por outro lado, se as cordas vocais estão alinhadas

e em contato ao longo de sua extensão, barrando assim a passagem do ar através da glote, a pressão que se forma por detrás dessa barreira avoluma-se de tal maneira que faz com que as duas dobras musculares sejam jogadas, abruptamente, para os lados em movimentos oscilatórios rapidíssimos, envolvendo em média de 60 a 500 oscilações por segundo. Acusticamente, esse movimento oscilatório é medido em termos de ciclos por segundo (cps) ou Hertz (Hz). Os sons que são produzidos com esse mecanismo são chamados de *sonoros*.

Contudo, como mencionamos anteriormente, nem só os sistemas respiratório e fonatório entram em cena na produção dos sons da fala. Acima da região glotal, na região conhecida como *supraglotal*, temos duas grandes câmaras de ressonância: a *cavidade oral* e a *cavidade nasal*. Isso quer dizer que o ar proveniente dos pulmões, e quase sempre transformado em energia vocal ao passar pela laringe, vai ter que encontrar, de alguma maneira, o seu caminho de saída. Existem, a esse respeito, três possibilidades:

- se o *véu palatino* estiver em sua posição mais frequente, *i.e., elevado*, e tocando a parede posterior da faringe, e, conseqüentemente, fechando a passagem para a cavidade nasal, não existe outra possibilidade de escape para o ar, senão que pela cavidade oral. Sons produzidos dessa forma são chamados de *orais*.
- se o *véu palatino* estiver *rebaixado*, o ar vai poder escapar pela cavidade nasal e pela cavidade oral ao mesmo tempo. Esse é o mecanismo de escape das *vogais nasalizadas*.
- contudo, existem, ainda, os casos em que o ar vai encontrar um *obstáculo* em algum ponto da *cavidade oral* impedindo sua passagem, mas vai conseguir sair pelo nariz devido à abertura da passagem que dá acesso à cavidade nasal, causada pelo *rebaixamento* do véu palatino. Esse é o mecanismo de escape que caracteriza a produção das *consoantes nasais*.

A *cavidade nasal* é basicamente composta pela faringe nasal, o espaço nasal e as fossas nasais, órgãos que, individualmente, não desempenham nenhum papel específico na produção dos diferentes sons da fala, a não ser o de funcionarem, conjuntamente, como uma grande caixa de ressonância, cuja forma e volume influenciam a qualidade da voz do indivíduo.

A *cavidade oral*, por outro lado, vai funcionar de forma bastante ativa durante a produção dos sons. Além de servir, de forma semelhante à cavidade nasal, como uma importante câmara de ressonância, as diferentes configurações dos órgãos que a compõem vão ser responsáveis pelas diferentes articulações ou gestos articulatórios que determinam qualidades sonoras distintas.

A cavidade oral tem sido dividida, para fins de estudo, em partes *móveis* e partes *estacionárias*. Os diferentes tipos de articulação, portanto, resultam das possíveis combinações das partes móveis com as estacionárias.

Dos órgãos móveis, o mais flexível é a *língua*, cujas diferentes partes vão se articular de tal forma a possibilitar uma grande variedade de sons, tanto em relação ao *ponto*, bem como ao *modo* ou *maneira de articulação*. As três partes da língua que desempenham papel relevante na produção da maior parte dos sons encontrados nas línguas naturais são: a *ponta* (ou *ápice*), a *lâmina* e o *dorso*. Os outros órgãos móveis da cavidade oral são os *lábios*, que, devido às possibilidades combinatórias, costuma-se dividir em *lábio superior* e *lábio inferior*.

As partes *estacionárias* da cavidade oral são, basicamente, a *arcada dentária* e as diferentes *partes do céu da boca*:

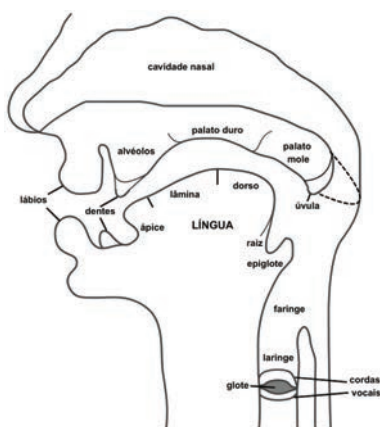
- *dentes superiores*;
- *dentes inferiores*;
- *alvéolos*: parte posterior e macia da gengiva superior, encontrada logo atrás e acima dos dentes superiores;
- *palato duro*: zona de consistência óssea localizada na parte anterior da abóbada palatina, em seguida aos alvéolos;
- *palato mole* (ou *véu palatino*): parte da abóbada palatina de consistência macia, encontrada em seguida ao palato duro;
- *úvula*: apêndice situado na extremidade do véu palatino, comumente conhecido como “campainha”. Pode adquirir movimento oscilatório a depender da força da camada de ar que se desloca pela região velo-faríngea, embora este não seja voluntário como no caso dos órgãos articulatórios, efetivamente móveis.

Além desses órgãos encontrados na cavidade oral propriamente dita, não podemos nos esquecer do longo canal que liga a laringe às cavidades oral e nasal: a *faringe*.

Em síntese, o ato de fala implica na integração dos três sistemas básicos: de *respiração* (que produz o ar), de *fonação* (que transforma o ar em energia sonora) e de *articulação* (que, através das diferentes configurações adotadas pelos diferentes órgãos da fala, determina a qualidade dos sons a serem produzidos).

Isso quer dizer que o ar que vem dos pulmões, a depender da posição das cordas vocais, transforma-se em energia na laringe e segue à procura de um caminho para escapar (*nasal, oral ou oral-nasal*). Contudo, antes de escapar, através de qualquer uma dessas três maneiras, o ar encontra configurações distintas dos diferentes órgãos da cavidade oral ou ao longo do tubo faríngeo. São exatamente essas configurações distintas que determinam os diferentes sons que produzimos. Em geral, essas diferentes configurações são expressas em termos dos *pontos* (ou *zonas*) e dos *modos de articulação*, no que diz respeito às articulações consonantais; e em termos da *altura* e do *avanço/recuo da língua*, em relação às vogais.

Figura 1 – O aparelho fonador humano



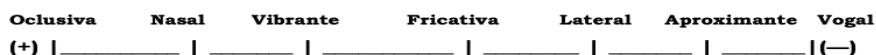
Fonte: elaborada pela autora<sup>3</sup>.

3 Nomenclatura/nomeação das tabelas, quadros, figuras e demais elementos foram mantidos com a formatação preferida pela autora.

## As consoantes

Os *Modos* referem-se aos tipos de obstáculos encontrados pelo ar durante sua trajetória de escape. Nesse sentido, podemos considerar a existência de uma escala relativa de *Estreitamento oral*, em cujos extremos teríamos o grau máximo de obstáculo para a passagem do ar (o que caracteriza a produção das *Oclusivas*) e o grau mínimo de impedimento, em que o ar se desloca pela cavidade oral livremente (marcando a produção das *Vogais*).

Figura 2 – Escala relativa de Estreitamento Oral



Fonte: elaborada pela autora.

Nas descrições articulatórias encontradas nos livros mais recentes de Fonética, a produção das consoantes é descrita como ocorrendo em três fases distintas: “Aproximação” (*approach*), “Manutenção” (*hold*) e “Soltura” (*release*) (Ladefoged, 1975).

A “Aproximação” refere-se ao momento em que um articulador móvel se aproxima de um outro articulador (móvel ou estacionário). Durante o momento da “Manutenção”, os órgãos articulatórios, após a aproximação, permanecem na postura adotada para tal gesto articulatório específico, mantendo contato firme, como ocorre no caso das oclusivas, por exemplo; ou apenas permanecem aproximados, deixando apenas um estreito canal para o escape do ar, como ocorre no caso das fricativas. A “Soltura”, então, refere-se ao afastamento dos órgãos articulatórios, após a fase de manutenção de um determinado gesto articulatório.

Assim, no extremo da escala que representa o grau máximo de estreitamento do canal oral, encontra-se a consoante *Oclusiva* (ou *Plosiva*, como é também conhecida). Nesse sentido, a *Oclusiva* figura como a consoante *mais consoante* que o aparato fonador humano é capaz de produzir. Na articulação da *Oclusiva*, são suas duas últimas fases, distintas, que lhe conferem sua característica primordial:

- *oclusão*: após a aproximação dos órgãos articulatórios, vai haver uma obstrução total e absoluta à passagem do ar em algum ponto da cavidade oral;
- *explosão*: a pressão de ar formada atrás da barreira articulatória cresce de tal forma em volume que separa, abruptamente, os articuladores responsáveis pela oclusão, criando um efeito semelhante ao de uma explosão, no momento da soltura. Essa sequência de movimentos articulatórios pode ser, facilmente, observada no caso das oclusivas bilabiais, [p] e [b], que nos fornecem boas pistas visuais.

É importante observar que essa articulação oclusiva, anteriormente referida, tem um escape de ar exclusivamente oral, visto que a passagem de ar que dá acesso à cavidade nasal encontra-se fechada durante toda sua produção, devido ao véu palatino encontrar-se em sua posição mais frequente, tocando a parede posterior da nasofaringe. Contudo, é possível haver uma articulação com bloqueio oral (isto é, em que exista *oclusão* na fase de manutenção), mas que, durante o momento da soltura, haja rebaixamento do véu palatino e a passagem que dá acesso à cavidade nasal fique, conseqüentemente, aberta. Dessa forma, o ar pulmonico encontra sua saída pelo trato nasal, antes mesmo que haja a explosão. Sons produzidos com esse tipo de mecanismo são chamados de consoantes oclusivas com escape nasal. Esse fato fonético, contudo, só é possível quando a saída nasal tem o mesmo ponto de articulação que a oclusiva. Casos como estes são bastante conhecidos na fonética da língua inglesa. Palavras como *kitten* (gatinho) e *button* (botão) são pronunciadas, por falantes nativos, como [ˈkɪtɪn] e [ˈbʌtɪn], respectivamente. Tantas são as possibilidades articulatórias permitidas pelo aparato fonador humano que, além de oclusivas com escape nasal, pode haver, também, oclusivas com escape lateral. De forma semelhante ao escape nasal, o escape lateral da oclusiva se dá quando a pressão de ar formada atrás da barreira de oclusão acha saída, através do rebaixamento das laterais da língua. Este é o caso da pronúncia das palavras inglesas *little* (pouco) e *saddle* (sela); respectivamente, [ˈlɪtl̩] e [ˈsædl̩].

As consoantes *Nasais* (também conhecidas por *Oclusivas nasais*, devido a sua semelhança às *Oclusivas*, anteriormente descritas) têm as fases de “Aproximação” e de “Manutenção” semelhantes às das *Oclusivas*, contudo seu momento de soltura se dá de forma diferenciada, uma vez que,

devido ao rebaixamento do véu palatino durante toda a articulação, não vai haver formação de pressão por trás do obstáculo oral, e o ar vai escapar sem maiores problemas pela cavidade nasal. Nessa classe, incluem-se sons como a bilabial [m] e a dento-alveolar [n] que ocorrem no português, por exemplo.

Em alguns casos, em vez de encontrar um obstáculo real em seu caminho, como no caso das *Oclusivas* e das *Nasais*, o ar pulmonico força passagem por um estreito canal oral (devido à aproximação de dois articuladores) ou faríngeo (devido a um tensionamento dos músculos da parede da faringe e da raiz da língua). Esses sons são chamados de consoantes *Fricativas*, devido ao fluxo turbulento de ar audível quando de sua produção, ou *Constritivas*, devido ao estreitamento ou constrição que se forma ao longo do canal orofaríngeo. Existe uma grande variedade de articulações fricativas, realizadas, basicamente, a partir de diferentes movimentos das diferentes partes da língua em direção a pontos distintos da abóbada palatina. São exatamente essas configurações da língua as responsáveis pela existência da grande variedade de sons fricativos encontrados nas línguas naturais, como [f], [v], [s], [z], [ʃ] e [ʒ] (todos eles encontrados no inventário de sons do português).

Durante a produção de alguns sons, o ar, ao forçar sua passagem pela cavidade oral, lança um articulador em movimento, gerando, assim, uma quantidade de vibrações compatível à intensidade da pressão do ar. Essas consoantes são chamadas de *Vibrantes*. Os órgãos articuladores que podem ser colocados em movimento oscilatório (ou vibratório) na produção dos sons da fala são, basicamente, dois: a *língua* (mais especificamente, seu *ápice* ou *ponta*) e a *úvula*. Os sons que, em sua produção, têm três ou mais movimentos oscilatórios são conhecidos como *Vibrantes múltiplas*. A pronúncia do “erre dobrado” de **CARRO**, em espanhol e em alguns dialetos do português, ilustra a vibrante múltipla lingual. Já o erre do francês, exemplifica a realização da vibrante múltipla uvular. Contudo, se o que ocorrer for apenas um movimento oscilatório leve, que joga um articulador (geralmente, a ponta da língua) contra outro (em geral, a região dento-alveolar ou pós-alveolar), temos, então, uma *Vibrante simples* – um *Flape* ou *Tape* (a depender da intensidade do deslocamento da língua) –, como ocorre na pronúncia do erre de **BARATA**.



Os sons *Laterais*, conforme indica o nome, originam-se de articulações em que vai haver um bloqueio à passagem do ar apenas na região central da cavidade oral, *i.e.*, a língua vai tocar algum ponto do céu da boca, mas o ar vai poder escapar, sem dificuldade, por um ou pelos dois lados da língua. Durante o momento de escape, dependendo do grau de estreitamento do canal lateral, pode haver uma saída turbulenta do ar, caso em que se obtém uma *Lateral fricativa* (que, em português, só ocorre como uma manifestação de distúrbio articulatorio). Quando o escape não é fricativo, então o fluir livre do ar mais se assemelha ao de uma aproximante, razão pela qual a lateral tem sido considerada, na Fonética mais atual, como uma *Aproximante lateral* (a esse respeito, ver Ladefoged, 1975, 1990, 1996, 2001 e I.P.A., 1989, 1999).

As *Aproximantes* (*não laterais*) são sons produzidos com um canal oral quase totalmente desimpedido, e, por essa razão, assemelham-se muito às vogais (motivo pelo qual alguns de seus elementos são, comumente, chamados de *Semivogais*). Durante sua produção, a aproximação de um articulador em direção a outro faz com que um pequeno estreitamento se forme no trato oral, sem, contudo, provocar turbulência durante o escape do ar. Entre as *Aproximantes*, sobressaem as conhecidas *Semivogais*, que são, articulatoriamente, definidas como *glides*, ou movimentos rápidos da língua, que podem adquirir duas direções: tendo como ponto de partida uma posição de vogal alta (anterior ou posterior) e se direcionando para outra posição vocálica de maior proeminência acústica, ou fazendo o percurso inverso. É, contudo, a natureza rápida e transitória desses sons, associada à força expiratória presente em sua produção, que os caracteriza como consonantais. Em geral, distinguem-se duas articulações semivocálicas:

- [j] – *Aproximante palatal*, em cuja articulação a lâmina da língua levanta-se em direção à área palatal, mas não chega a formar um canal tão estreito nesse ponto, como ocorre com a consoante fricativa [ç]. Na literatura específica de fonética e fonologia do português, tem sido comum o uso do símbolo [y] para designar esse som; e, por coerência a essa tradição, embora estejamos, sempre, nos apoiando na notação fornecida pelo Alfabeto Internacional de Fonética (IPA), estamos, aqui, preferindo utilizar o [y] por questões meramente práticas;

- [w] – *Aproximante labial-velar*, produzida através da aproximação simultânea dos lábios e da parte posterior da língua em direção ao véu palatino.

Note-se que em alguns casos, como ocorre em português, as *Semivogais* são de natureza mais vocálica que consonantal, *i.e.*, embora se caracterizem por movimentos rápidos, o grau de abertura do canal de escape do ar assemelha-se ao das vogais altas (Mateus, 1976).

Partindo desse último tipo de articulação, chegamos, então, ao grau máximo de abertura do canal oral – que ocorre na produção das *Vogais*, que são, a seguir, tratadas em maior detalhe.

A partir dessa abordagem “escalar” dos modos de articulação, consegue-se melhor entender e apreciar as relações existentes entre consoantes e vogais – que deixam de ser encaradas como dois tipos absolutamente distintos e mutuamente exclusivos de segmentos sonoros e passam a ser vistas como graus relativos em uma escala de fechamento/abertura do trato oral. Assim, pode-se explicar de forma mais convincente (do que nas abordagens fonéticas mais tradicionais) os processos, atestados em diversas línguas, de consoantes que se enfraquecem e viram vogais. No português, esse processo pode ser ilustrado pela semivocalização da lateral pós-vocálica que ocorre em formas como **SAL** [saw], **SOLTO** [sowtu].

Finalmente, para que as articulações consonantais possam ser plenamente descritas e apreciadas, uma vez que já examinamos os tipos de obstáculos que impedem, em maior ou menor grau, o fluxo do ar (ou os *Modos de articulação*), falta-nos, ainda, investigar em que *pontos* ou *zonas* do trato oral tais impedimentos vêm a ocorrer. Mais precisamente, resta-nos descrever os *Pontos de articulação*.

Os *Pontos* ou *Zonas de articulação* são determinados a partir dos movimentos dos órgãos móveis da fala em relação aos órgãos estacionários. Em geral, são descritos a partir da região mais anterior do aparelho fonador (ou seja, daquela mais visível ao olho humano):

- sons produzidos com a participação dos lábios superior e anterior são chamados de *Bilabiais*, por exemplo, [m] e [b], como nas palavras **MALA** e **BALA**;

- sons produzidos através de contato ou aproximação entre o lábio inferior e a arcada dentária superior são chamados de *Labiodentais*, por exemplo, [f] e [v], como nas palavras **F**ALA e **V**ALA;
- sons produzidos através de contato ou aproximação da ponta da língua com a arcada dentária superior (e muitas vezes também com os dentes inferiores) são chamados de *Dentais*, por exemplo, [θ] e [ð], enquanto consoantes iniciais nas palavras inglesas **TH**INK e **TH**IS. Em português, esses sons podem ocorrer como pronúncias articulatoriamente imprecisas de [s] e [z], respectivamente;
- sons produzidos com contato ou aproximação entre a ponta da língua (e/ou parte de sua lâmina) e os alvéolos são classificados como *Alveolares*. No Português Brasileiro (PB), o mais sensato é falar-se em uma área *Dento-Alveolar*, em que diversos segmentos podem ser realizados, tanto de forma mais *dentalizada* (a depender da vogal imediatamente subsequente e também do dialeto geográfico) como *alveolarizada*, como ocorre, por exemplo, nas pronúncias de [t], [n], [s] e [z] nas palavras **T**ELA, **N**ELA, **S**ELA e **Z**ELA, devido ao fenómeno da *Coarticulação*, visto que a vogal [ɛ] (assim como [e] e, por excelência, [i]) articula-se na área palatal, o que faz com que, naturalmente, a consoante precedente seja produzida um pouco mais para trás;
- sons produzidos com aproximação da lâmina da língua à região do palato duro (e áreas bastante contíguas) são chamados de *Palatais*, como [ʎ], [ɲ] e [ʃ] na terceira sílaba das palavras **ENCALH**ADO, **ACA-NH**ADO e **ENCAIX**ADO;
- alguns sons são articulados através da aproximação ou contato entre o dorso da língua e o véu palatino. Estes são os sons *Velares*, como o [k] e o [g] em **C**ALO e **G**ALO;
- existem, também, sons produzidos por um contato (caso das oclusivas uvulares do esquimó) ou aproximação da parte posterior do dorso da língua com a úvula ou por sua oscilação devido à forte pressão do ar pulmônico, conforme explicitado anteriormente. Um exemplo de realização *Uvular* é a pronúncia do erre vibrante, [ʀ], em alguns dialetos do português, em particular do Português Europeu (PE), ou a fricativa [ʁ] em alguns dialetos do PB, como na palavra **R**ATO, [ʀatu] e [ʁatu] respectivamente (Mateus, 1990);



A autora é licenciada em Português e Inglês pela Pontifícia Universidade Católica do Rio de Janeiro (PUC-RJ), obteve o M. A. em Linguística na Universidade do Kansas, o M. Phil. e o Ph. D em Fonética e Linguística na Universidade de Londres. Fez pós-doutoramento na Universidade do Texas. É professora titular da Universidade Federal da Bahia (UFBA) e pesquisadora em: aquisição da fonologia e do léxico do português em condições típicas e não típicas, acessibilidade e estudos surdos. Coordenou o Polo UFBA do curso de licenciatura e bacharelado em Letras Libras (2006-2012), o Polo UFBA do curso de graduação em Pedagogia bilíngue EAD (2017-2023). Coordena o Núcleo de Apoio à Inclusão da Pessoa com Necessidades Especiais (NAPE/ PROAE/UFBA) (2015-).



Esta obra apresenta os aspectos fonoarticulatórios de composição e de combinação do inventário de sons do português, como também os padrões de simplificação do sistema fonológico durante o desenvolvimento da fala e da escrita em situações típicas e não típicas. Além disso, traça um interessante paralelo entre padrões aquisicionais e padrões de simplificação fonológica marcadores de estigmatização sociolinguística. Aborda, ainda, a contagem de frequência de ocorrência de padrões segmentais, silábicos, lexicais, intrassilábicos e intersilábicos, mostrando, estatisticamente, os padrões que têm maior representatividade na língua, e são mais simplificados na fala infantil e na fala adulta, com destaque especial para os Erros de Ordenação Serial, e a Permutação Segmental Intersilábica da fala infantil.

**fapesb**  
Fundação de Amparo  
à Pesquisa do Estado da Bahia



GOVERNO DO ESTADO  
**BAHIA**  
SECRETARIA DE CIÊNCIA,  
TECNOLOGIA E INOVAÇÃO

ISBN 978-65-5630-829-6



9 786556 308296